

**PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK MENINGKATKAN AKURASI
SENTIMEN ANALISA DENGAN MENGGUNAKAN ALGORITMA SIMILARITAS WU
DAN PALMER**



**Disusun sebagai salah satu syarat menyelesaikan Program Studi Strata I pada Jurusan
Informatika Fakultas Komunikasi dan Informatika**

Oleh:

Firnadi Rio

L 200 130 108

**PROGRAM STUDI INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

2017

HALAMAN PERSETUJUAN

**PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK
MENINGKATKAN AKURASI SENTIMEN ANALISA DENGAN
MENGUNAKAN ALGORITMA SIMILARITAS WU DAN PALMER**

PUBLIKASI ILMIAH

Telah dipertahankan di depan Dewan Pengaji
Fakultas Komunikasi dan Informatika
Universitas Muhammadiyah Surakarta
Pada hari Kamis, 03 Agustus 2017
dan dinyatakan oleh:

oleh:

Dewan Pengaji:

1. Endang Wahyu Pamungkas, S.Kom, M.Kom.

(Ketua Dewan Pengaji)

2. Yodik Indra Kurnia, S.T, M.Eng.

(Anggota I Dewan Pengaji)

3. Dika Triyowati, S.T, M.Eng.

(Anggota II Dewan Pengaji)

FIRNADI RIO

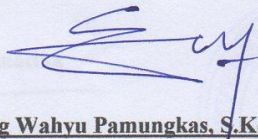
L 200 130 108

Telah diperiksa dan disetujui untuk diuji oleh:

Tanggal 03 Agustus 2017

Dipertahankan,

Dosen Pembimbing



Endang Wahyu Pamungkas, S.Kom, M.Kom.

NIK. 100.1704

HALAMAN PENGESAHAN

**PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK
MENINGKATKAN AKURASI SENTIMEN ANALISA DENGAN
MENGUNAKAN ALGORITMA SIMILARITAS WU DAN PALMER**

OLEH

FIRNADI RIO

L 200 130 108

**Telah dipertahankan di depan Dewan Penguji
Fakultas Komunikasi dan Informatika
Universitas Muhammadiyah Surakarta
Pada hari Kamis, 03 Agustus 2017
dan dinyatakan telah memenuhi syarat**

Dewan Penguji:

1. Endang Wahyu Pamungkas, S.Kom., M.Kom.

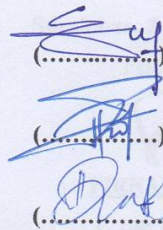
(Ketua Dewan Penguji)

2. Yogiek Indra Kurniawan, S.T., M.T.

(Anggota I Dewan Penguji)

3. Dyah Priyawati, S.T., M.Eng.

(Anggota II Dewan Penguji)


(.....)
(.....)
(.....)

Publikasi ilmiah ini telah diterima sebagai salah satu persyaratan

Untuk memperoleh gelar sarjana


Tanggal 03 Agustus 2017

Mengetahui,

**Dekan
Fakultas Komunikasi dan Informatika**


Nurgiyatna, S.T., M.Sc., Ph.D.
NIK. 881

**Ketua Program Studi
Informatika**


Dr. Heru Supriyono, M.Sc.
NIK: 970

PERNYATAAN

Jl. A Yani Tremol Pos 1 Paingan 57102 Surakarta 27113717417, 719483 Fax (0271) 714443
Surakarta 27102 Indonesia Web: <http://informatika.ums.ac.id> Email: informatika@ums.ac.id

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam naskah dan disebutkan dalam daftar pustaka.

Apabila kelak terbukti ada ketidakbenaran dalam pernyataan saya di atas, maka akan saya pertanggungjawabkan sepenuhnya.

FIRNADI RIO

L200130108

PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK
MENINGKATKAN AKURASI SENTIMEN ANALISA DENGAN
MENGUNAKAN ALGORITMA SIMILARITAS WU DAN
PALMER

Surakarta, 3 Agustus 2017

Penulis



FIRNADI RIO

L 200 130 108



**UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI INFORMATIKA**

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id>. Email: informatika@ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

246 /A.3-IL.3/INF-FKI/VIII/2017

Assalamu'alaikum Wr. Wb

Biro Tugas Akhir Program Studi Informatika menerangkan bahwa :

Nama : FIRNADI RIO
NIM : L200130108
Judul : PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK
MENINGKATKAN AKURASI SENTIMEN ANALISA DENGAN
MENGUNAKAN ALGORITMA SIMILARITAS WU DAN
PALMER
Program Studi : Informatika
Status : **Lulus**

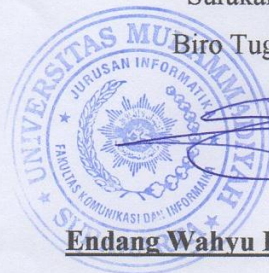
Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Tugas Akhir,
dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 8 Agustus 2017

Biro Tugas Akhir Informatika



Endang Wahyu Pamungkas, S.Kom., M.Kom.

Feedback Studio - Google Chrome

Aman

https://ev.turnitin.com/app/carta/en_us/?s=1&lang=en_us&u=1057550080&o=835767123

turnitin

13 of 27

?

Match Overview

×

20

20

Submitted to Universita...

Student Paper

12% >

eprints.ums.ac.id

Internet Source

4% >

Submitted to National I...

Student Paper

1% >

www.ijer.in

Internet Source

1% >

www.sepin.org

Internet Source

1% >

bmcbioinformatics.bio...

Internet Source

1% >

englishfirst.co.id

Internet Source

<1% >

pt.scribd.com

Internet Source

<1% >

blog.ephi.web.id

Internet Source

<1% >

www.insing.com

Internet Source

<1% >

es.scribd.com

Internet Source

<1% >

Activate Windows

Go to PC settings to activate Windows.

1

Abstrak

Saat ini hampir semua kebutuhan informasi manusia dapat didapatkan hanya dengan mengakses halaman *web* tertentu. Bahkan hampir semua kebutuhan sehari-hari juga dapat didapatkan dengan memanfaatkan berbagai teknologi digital yang ada. Dengan semakin banyaknya pengguna layanan di internet, maka semakin banyak pula informasi yang dapat dihimpun. Dari informasi tersebut kita dapat mengidentifikasi atau menganalisa kata kalimat yang baik atau kurang baik. Untuk melakukannya dibutuhkan sentimen analisis untuk melakukan pengidentifikasian dari kalimat-kalimat tersebut. Dari informasi-informasi tersebut banyak kalimat yang memiliki sifat ambigu. Ambigu merupakan fenomena dalam penggunaan bahasa yang merujuk kepada suatu atau ungkapan yang memiliki makna lebih dari satu. Terlepas dari kenyataan bahwa ambiguitas dalam bahasa merupakan bagian penting dari bahasa, sering menjadi kendala atau masalah untuk seseorang memahami suatu kalimat. Untuk bisa menghasilkan makna yang tepat untuk kalimat ambigu tersebut, dibuat penelitian menggunakan metode algoritma similariti Wu & Palmer. Dengan menggunakan metode ini kita memperoleh peningkatan dari nilai *precision*, *recall*, dan *accuracy*. Dari nilai *precision* mengalami peningkatan sebesar 0,06 samapi 0,08. Nilai *recall* mengalami peningkatan sebesar 0,22 sampai 0,24. Dan nilai *accuracy* mengalai peningkatan sebesar 0,04 samapai 0,06. Dapat dilihat dari nilai peningkatan yang terjadi maka motode *wu & palmer* lebih efektif untuk digunakan. Dengan penelitian ini diharapkan dapat membantu penelitian khususnya berkaitan

Page: 1 of 12

Word Count: 3228

Q

Q

PENDETEKSIAN AMBIGUITAS MAKNA KATA UNTUK MENINGKATKAN AKURASI SENTIMEN ANALISA DENGAN MENGGUNAKAN ALGORITMA SIMILARITAS WU DAN PALMER

Abstrak

Saat ini hampir semua kebutuhan informasi manusia dapat didapatkan hanya dengan mengakses halaman *web* tertentu. Bahkan hampir semua kebutuhan sehari-hari juga dapat didapatkan dengan memanfaatkan berbagai teknologi digital yang ada. Dengan semakin banyaknya pengguna layanan di internet, maka semakin banyak pula informasi yang dapat dihimpun. Dari informasi tersebut kita dapat mengidentifikasi atau menganalisa kata kalimat yang baik atau kurang baik. Untuk melakukannya dibutuhkan sentimen analisis untuk melakukan pengidentifikasian dari kalimat-kalimat tersebut. Dari informasi-informasi tersebut banyak kalimat yang memiliki sifat ambigu. Ambigu merupakan fenomena dalam penggunaan bahasa yang merujuk kepada suatu atau ungkapan yang memiliki makna lebih dari satu. Terlepas dari kenyataan bahwa ambiguitas dalam bahasa merupakan bagian penting dari bahasa, sering menjadi kendala atau masalah untuk seseorang memahami suatu kalimat. Untuk bisa menghasilkan makna yang tepat untuk kalimat ambigu tersebut, dibuat penelitian menggunakan metode algoritma similariti Wu & Palmer. Dengan menggunakan metode ini kita memperoleh peningkatan dari nilai *precision*, *recall*, dan *accuracy*. Dari nilai *precision* mengalami peningkatan sebesar 0,06 samapi 0,08. Nilai *recall* mengalami peningkatan sebesar 0,22 sampai 0,24. Dan nilai *accuracy* mengalai peningkatan sebesar 0,04 samapai 0,06. Dapat dilihat dari nilai peningkatan yang terjadi maka motode *wu & palmer* lebih efektif untuk digunakan. Dengan penelitian ini diharapkan dapat membantu penelitian khususnya berkaitan dengan analisis sentiment untuk kalimat ambiguitas dengan metode Wu & Palmer.

Kata Kunci : sentiment analysis, SentiWordNet, Wu & Palmer

Abstract

Nowadays, almost all the needs of human information can be obtained only by accessing web pages. In fact, almost all daily necessities can also be obtained by utilizing various digital technologies . Increasing of internet user , had effect a lot of information can be collected. From this information we can identify or analyze word sentence is good or bad. In this case sentiment analysis is needed to identification of these sentences. More the sentences of information have ambiguous. Ambiguous is a phenomenon in the use of language refers to a phrase more than one. Despite the fact that the ambiguity in the language is an important part of the language, is often an obstacle for a human to understand a sentence. In this case, to produce the appropriate meaning of the sentence ambiguous. Using this method we get an increase of precision, recall, and accuracy values. Of the precision values experienced increased by 0,06 until 0,08. Recall value increased by 0,22 until 0,24. And the accuracy value increased by 0,04 until 0,06. Can be used from the increased value that occurs then method Wu & Palmer is more effective to use. Make research using the algorithm method similarity Wu & Palmer. In this case the research is expected to aid in particular regard to sentiment analysis on ambiguity sentence by Wu & Palmer method.

Keywords: sentiment analysis, SentiWordNet, Wu & Palmer

1. PENDAHULUAN

Sekarang perkembangan dunia digital semakin meningkat pesat. Sehingga pemanfaatan internet di Indonesia khususnya mulai meningkat drastis akhir-akhir ini. Hal tersebut memicu pertumbuhan dan perkembangan teknologi World Wide Web. Sekarang ini hampir setiap informasi untuk memenuhi kebutuhan manusia mudah didapatkan hanya dengan mengakses halaman *web* tertentu. Bahkan hampir semua kebutuhan sehari-hari juga bisa didapatkan dengan memanfaatkan berbagai teknologi digital yang ada.

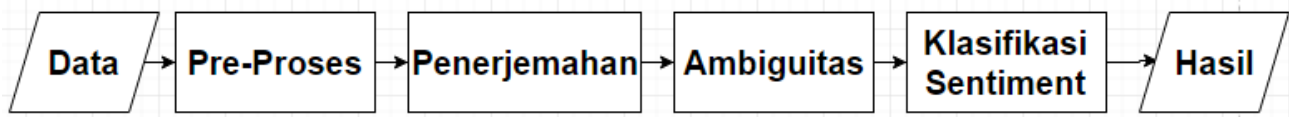
Dengan semakin banyaknya pengguna layanan di internet, maka semakin banyak pula informasi yang dapat dihimpun. Dari informasi-informasi tersebut banyak kalimat yang memiliki sifat ambigu. Ambigu merupakan fenomena dalam penggunaan bahasa yang merujuk kepada suatu atau ungkapan yang memiliki makna lebih dari satu. Ada juga yang mengatakan ambiguitas adalah kemungkinan menafsirkan sebuah ungkapan dalam dua atau lebih cara yang berbeda (Kapadia, D., and Jufrizal, 2013). Dengan kata lain, sesuatu yang ambigu ketika itu bisa dipahami dalam dua atau lebih indera dan cara. Terlepas dari kenyataan bahwa ambiguitas dalam bahasa merupakan bagian penting dari bahasa, sering menjadi kendala atau masalah untuk seseorang memahami suatu kalimat.

Sudah banyak penelitian yang mencoba untuk menyelesaikan permasalahan yang berkaitan dengan *sentiment analysis*. *Sentiment analysis* adalah jenis pemrosesan Bahasa alami untuk mendeteksi opini masyarakat tentang topik tertentu (Vinodhini, G. And RM. Chandrasekaran, 2012). Ada juga yang berpendapat bahwa analisa sentimen sering juga dikenal dengan *opinion mining* adalah studi komputasi dari pendapat, sentimen, sikap, serta emosi yang disajikan dalam sebuah teks (Liu, B., 2012).

Salah satu penelitian sebelumnya mencoba untuk melakukan *sentiment analysis* multilingual (Denecke, K., 2008). Dalam penelitian tersebut digunakan PROMT sebagai *translator* dan juga memanfaatkan SentiWordNet sebagai *dictionary*. Namun pada saat itu SentiWordNet yang digunakan masih versi 1.0. Hasil dari penelitian tersebut masih kurang maksimal dengan segala keterbatasan teknologi yang ada. Salah satunya yaitu ketergantungan hasil dengan kualitas mesin *translator* yang digunakan saat itu. Maka dari itu, pada penelitian ini akan digunakan translator yang memiliki teknologi yang lebih mutakhir dan banyak digunakan yaitu Bing Translator dan Google Translator. Pada penelitian yang lain menggunakan metode *First Sense* hasilnya dirasa kurang baik, karena terjadi ketidaksesuaian antara klasifikasi sentimen manual dan hasil dari sistem (Kusumawati, I., & Pamungkas, E. W., 2017).

Dengan timbulnya masalah-masalah diatas, disini akan dilakukan pembuatan sistem untuk membantu menentukan makna kata yang tepat untuk suatu kalimat yang ambigu menggunakan metode *sentimen analisa*, dengan algoritma similariti Wu & Palmer. Sistem tersebut akan memberikan makna terbaik untuk kalimat yang mengandung kata ambigu. Sehingga dengan penelitian ini diharapkan dapat membuka peluang penelitian yang lebih baik kedepannya.

2. METODOLOGI



Gambar 1. Metode Penelitian

Gambar 1 di atas menunjukkan alur dari jalannya metode yang akan dilakukan dalam penelitian ini. Data yang menjadi masukan adalah berupa kalimat. Kalimat yang dimasukan akan langsung dilakukan proses translasi ke Bahasa Inggris dengan menggunakan translator yang tersedia. Setelah kalimat diterjemahkan kemudian dilanjutkan dengan proses analisis sentiment dengan memanfaatkan Wu & Palmer SentiWordNet. Keluaran dari serangkain proses ini adalah berupa nilai sentiment dari kalimat yang menjadi masukan dan makna kata terbaik dari kalimat yang dimasukkan. Untuk lebih jelasnya, berikut penjelasan yang lebih detail dari masing-masing proses yang ada pada diagram di atas.

2.1. Data

Data yang akan dipakai disini adalah data yang sudah diambil dari opini masyarakat yang menggunakan media social *blackberry messenger* (BBM), LINE, dan Whatsapp. Nantinya data ini yang akan dipakai untuk uji coba sistem tersebut sistem tersebut. Dalam bentuk excel.

2.2. Pre-Proses

Disini adalah proses dimana akan diperiksanya kata yang akan dimasukkan dalam sitem. Nantinya akan dilakukan pendeteksian kata yang tidak sesuai. Contoh kata „Nooooo“ akan di ubah menjadi kata „No“. Namun proses ini dilakukan secara manual.

2.3. Penerjemahan

Selanjutnya, ketika kata sudah benar, maka disini dilakukan proses penerjemahan dari suatu bahasa akan diterjemahkan kedalam Bahasa Inggris. Tool yang telah tersedia dan banyak diunakan akan dimanfaatkan untuk melakukan proses penerjemahan. Hasil dari proses terjemahan tadi tidak perlu lagi dikoreksi oleh manusia karena sudah dianggap benar. Google translator adalah tools translator yang akan digunakan untuk melakukan penerjemahan. Google translator tidak kalah dibandingkan Bing translator, namun sama halnya dengan Bing, Google tranlatot harus melakukan pengecekan manual untuk hasil yang berkualitas. Hasil dari Google translator tersebut akan langsung diprosen didalam program.

2.4. Proses Ambiguitas

Setelah kalimat menjadi berbahasa inggris, selanjutnya kalimat yang ada tadi akan dilakukan proses POS Tagging untuk mencari *part of speech* (POS) yang meliputi kata benda (*noun*), kata kerja (*verb*), kata sifat (*adjective*), dan kata keterangan (*adverb*) yang sesuai dengan konteks kalimat. *Part of speech* ini akan menentukan proses pencarian di dictionary SentiWordNet. Proses POS Tagging ini juga memanfaatkan library yang sudah ada banyak dipakai yaitu Stanford POS Tagger.

Proses POS Tagging dilakukan dengan menggunakan Stanford POS Tagger. Dari dua tahap ini akan dihasilkan daftar kata penyusun kalimat, lengkap dengan *Parts of Speech*. Sebagai contoh :

“best way to communicate with friends and colleagues.”

Hasil POS Tagger : best_JJS way_NN to_TO communicate_VB with_IN friends_NNS and_CC colleagues_NNS

Dari hasil POS Tagger diatas, hanya kata benda (*noun*), kata kerja (*verb*), dan kata sifat (*adjective*) yang akan digunakan dalam proses ini.

NN dikategorikan sebagai kata benda (*noun*), VB dikategorikan sebagai kata kerja (*verb*), JJ dikategorikan sebagai kata sifat (*adjective*).

Hasil proses ini akan digunakan pada tahap selanjutnya yaitu penghilangan ambiguitas.

Contoh kata yang dapat bernilai positif atau negatif yang mempengaruhi perhitungan nilai sentiment total, sebagai berikut:

“very clear sound and easy to use”

Dalam perhitungan dengan metode *Wu & Palmer* hasil dari kalimat tersebut adalah positif, yaitu POS Tagger = „clear_JJ, sound_NN, easy_JJ, use_VB

Tabel 1. Hasil Proses

ID Synset	Positif	Negatif
00428404 #clear	0,5	0
062781360 #sound	0	0
00749230 #easy	0,625	0,25
01158872 #use	0	0
Final Score	0,875	

Kata “clear” dengan POS *adjective* dalam sentiWordnet memiliki akumulasi skor positif dan negatif, salah satu synset “clear” dengan ID 00428404 memiliki skor positif 0,5 dan skor negatif 0 sehingga menghasilkan final score 0,875 berarti skor tersebut masuk dalam data positif. Synset “clear” dengan ID 03040836 memiliki skor positif 0 dan memiliki skor negatif 0,5 sehingga menghasilkan total skor negatif. Pada contoh kalimat diatas, metode *Wu & Palmer* memilih synset “clear” dengan ID 00428404 yang total skor positif nya menghasilkan nilai sentiment positif (0,875) sehingga sesuai dengan opini yang ditentukan. Jika dipilih synset “clear” dengan ID 03040836 maka total nilai sentiment nya adalah negatif (-0.5). sehingga tidak sesuai dengan opini yang ditentukan.

Setelah semua kata telah melakukan POS Tagger, proses selanjutnya adalah mencari makna kata yan paling tepat atau penghilangan ambiguitas. Algoritma yang digunakan adalah Wu Palmer dari Word Similarity for Java (WS4J) dan Wordnet. Berikut persamaan yang digunakan:

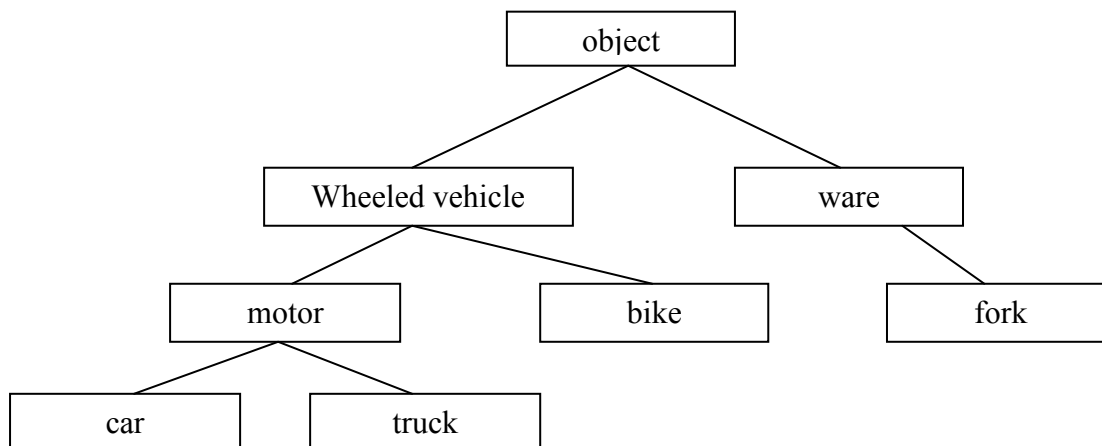
$$WUP(kata1, kata2) = -\log (2 \times DepthLCS / (Depth1 + Depth2)$$

DepthLCS merupakan titik temu antara dua kata yang ditinjau pada hirarki WordNet.

Depth1 merupakan kedalaman terpendek pada hirarki WordNet.

Depth2 merupakan kedalaman terpendek pada hirarki WordNet

Contoh perhitungan nya :



Gambar 2. Akar Perhitungan

Gambar 2 di atas akan dihitung synset dari truck dan bike. Maka kita tentukan dulu depthlcs nya, depthlcs berada pada kata wheeled vehicle dengan nilai 1. Depth1 terletak pada kata truck yang memiliki nilai 3, angka tersebut di ambil dari jarak antara truck dengan object. Depth2 terletak pada kata bike yang memiliki nilai 2, angka tersebut didapatkan sama dengan mendapatkan depth1 yaitu menghitung jarak antara bike dengan object. Setelah mendapat angka-angka tersebut maka akan dimasukkan kedalam persamaan (1).

$$WUP = -\log (2 \times 1 / (3 + 2)) = 0,301 \dots\dots\dots(1)$$

Setelah angka-angka tersebut dimasukkan dalam rumus maka kita nilai 0,301 dari 2 kata yang proses dalam rumus tersebut. Disini nanti dari sekian makna kata yang ada akan saling di terkaitkan dan akan di ambil nilai yang paling tinggi sebagai hasil makna kata nya. (Wicaksana, I. W. S., & Reza A. H., 2006). Dari proses yang dilakukan ini kita dapat menghasilkan makna kata paling tepat yang akan di proses di tahan selanjutnya yaitu Sentiment Clasification.

2.5. Proses Sentiment Clasification dengan WordNet

Pada bagian ini akan dilakukan proses klasifikasi sentiment per kata dalam kalimat serta perhitungan nilai sentiment total pada kalimat. Dari makna kata yang dihasilkan tadi maka akan memiliki nilai positif dan negatif yang diambil dari database sentiwordnet.

Rumus yang digunakan :

$$S_{positive} = \sum_{i \in T} positive\ score_i \dots\dots\dots (2)$$

$$S_{negative} = \sum_{i \in T} negative\ score_i \dots\dots\dots (3)$$

Nilai-nilai tersebut nantinya akan dijumlahkan positif di totalkan dan negatif juga ditotalkan selanjutnya nilai total positif akan dikurangi nilai total negatif menggunakan persamaan (2) dan (3) (Pamungkas, E. W., & Putri, D. P., 2016). Kemudian dari proses tersebut kita dapat mengetahui nilai dari makna yang tepat tersebut positif atau negatif, Kemudian dengan rumus dibawah ini kita menentukan orientasi semantiknya.

$$Sentence_{sentiment} \begin{cases} positive\ if\ S_{positive} > S_{negative} \\ neutral\ if\ S_{positive} = S_{negative} \\ negative\ if\ S_{positive} < S_{negative} \end{cases} \dots\dots\dots (4)$$

Dengan melihat persamaan di atas maka kita dapat menentukan sentiment dari suatu kalimat dengan cara membandingkan jumlah skor positif dan negatif. Jika skor positif lebih besar dari negatif maka hasilnya positif. Jika skor positif lebih kecil dari skor negatif maka hasilnya negatif. Dan jika skor positif sama dengan skor negatif maka hasilnya netral, persamaan (4) (Pamungkas, E. W., & Putri, D. P., 2016).

3. HASIL DAN PEMBAHASAN

Didalam pembahasan ini, nantinya akan dijelaskan mengenai bagai mana proses berjalan dan perbandingannya. Untuk datanya sendiri menggunakan dataset yang di ambil dari opini masyarakat yang menggunakan aplikasi social media seperti *Blackbarry Message* (BBM), Line, dan Whatsapp. Dari opini-opini yang ada terkumpul sebanyak 334 data, yang terdiri dari 131 opini positif, kemudian 203 opini negatif, yang diklasifikasi dengan cara manual.

Untuk membuktikan peningkatan performa dalam analisa sentimen dapat dilakukan dengan cara membandingkan teknik *Lexicon Based* menggunakan metode *Wu Palmer* dengan metode *First Sense*. Metode *Lexicon Based* mampu mengatasi dokumen yang sama tetapi memiliki sentimen berbeda dalam konteks yang berbeda (Pamungkas, 2016). Table 2 di bawah dapat menunjukkan peningkatan peforma yang terjadi.

Perbandiangan pengujian antara metode *Wu & Palmer* dengan metode *first sense* dan metode *average*. Untuk melihat peningkatan performa maka dilakukanlah pengujian tersebut dengan mengkategorikan dalam tiga kategori yaitu *precision*, *recall*, dan *accuracy*. *Precision* sendiri memiliki peran sebagai tingkatan keakuratan antara data klasifikasi secara manual dan keluaran yang dihasilkan oleh system. Kemudian *Recall*, disini *Recall* memiliki peran sebagai tingkatan kesuksesan system saat mendapatkan kembali data yang diuji. Dan yang terakhir adalah *Accuracy*, yang berguna sebagai alat ukur tingkatan kedekatan antara perkiraan dengan nilai yang sesungguhnya.

Untuk mendapatkan nilai-nilai tersebut menggunakan rumus perhitungan sebagai berikut:

3.1. *Precision* : $TP / (TP + FP)$

3.2. *Recall* : $TP / (TP + FN)$

3.3. *Accuracy* : $(TP + TN) / (TP + TN + FP + FN)$

Keterangan :

TP (*True Positif*) : Kondisi dimana manual positif atau manual netral terdeteksi positif atau netral .

TN (*True Negatif*) : Kondisi dimana manual negatif terdeteksi negatif.

FP (*False Positif*) : Kondisi dimana manual negatif terdeteksi positif atau netral.

FN (*False Negatif*) : Kondisi dimana manual positif atau manual netral terdeteksi negatif.

Tabel 2. Hasil Pengujian

PARAMETER	METODE		
	WU PALMER	FIRST SENSE	AVERAGE
PRECISION	0.49	0.41	0.43
RECALL	0.89	0.67	0.65
ACCURACY	0.59	0.53	0.55

Perbandiangan pengujian antara metode *Wu & Palmer* dengan metode *first sense* dan metode *average*. Untuk melihat peningkatan performa maka dilakukanlah pengujian tersebut dengan

mengkategorikan dalam tiga kategori yaitu *precision*, *recall*, dan *accuracy*. *Precision* sendiri memiliki peran sebagai tingkatan keakuratan antara data klasifikasi secara manual dan keluaran yang dihasilkan oleh system. Kemudian *Recall*, disini *Recall* memiliki peran sebagai tingkatan kesuksesan system saat mendapatkan kembali data yang diuji. Dan yang terakhir adalah *Accuracy*, yang berguna sebagai alat ukur tingkatan kedekatan antara perkiraan dengan nilai yang sesungguhnya.

Dari hasil tabel di atas dapat kita lihat perbandingan yang cukup jelas, pertama dapat dilihat dari nilai *precision* pada metode *first sense* yang memiliki nilai 0,41. Kemudian pada metode *average* memiliki nilai 0,43. Dan yang terakhir adalah metode *Wu & Palmer* yang memiliki nilai 0,49. Nilai tersebut adalah nilai paling tinggi di bandingkan dengan metode *first sense* dan metode *average*. Yang kedua dilihat dari nilai *recall*, pada metode *first sense* memiliki nilai 0,67. Kemudian pada metode *average* memiliki nilai 0.65. Dan kemudian metode yang terakhir adalah *Wu & Palmer* yang mengalami peningkatan cukup signifikan, disini nilai yang didapatkan mencapai 0,89. Pada perbandingan *recall* ini metode *Wu & Palmer* juga memiliki nilai yang paling tinggi dibandingkan nilai dari metode *first sense* dan metode *average*. Dan yang ketiga dapat kita lihat dari nilai *accuracy*, pada metode *first sense* memiliki nilai 0,53. Kemudian pada metode *average* memiliki nilai 0,55. Dan metode *Wu & Palemer* yang memiliki nilai tertinggi yaitu 0,59. Dari nilai-nilai tersebut dapat kita pahami bahwa nilai *precision* dan *accuracy* mengalami peningkatan yang hampir sama, dapat dilihat pada tabel peningkatan nilai *precision* hanya sekitar 0.06 sampai 0.08 saja, kemudian untuk nilai *accuracy* hanya mengalami peningkatan sebesar 0,04 sampai 0,06 saja, dan untuk nilai *recall* menunjukkan peningkatan yang cukup besar, dari nilai tersebut membuktikan algoritma similarity menggunakan metode *Wu & Palmer* dapat meningkatkan performa sistem dalam mendapatkan kembali data yang sedang diuji. Dapat dilihat nilai *precision* cukup rendah, hal tersebut di pengaruhi oleh data-data yang memiliki opini negatif pada perhitungan manual dan menghasilkan nilai positif pada program yang dijalankan. Contoh nya seperti kalimat “*was good at first, now just annoying with ads, feel slow and laggy. use high ram. drains battery*”, kalimat ini mengandung nilai negatif pada penghitungan manual dan positif pada penghitungan dengan system.

Tabel 3. Proses POS Tagging

WORD	POS
Was	VBD
Good	JJ
At	IN
First	RB
Now	RB
Just	RB
Annoying	VGB
With	IN
Ads	NNS
Feel	VBP
Slow	JJ
And	CC
Laggy	NN
Use	NN
High	JJ
Ram	NN
Drains	VBZ
Battery	NN

Tabel 4. Hasil sentiment

ID Synset	Skor Positif	Skor Negatif
02480111	0,75	0
00875650	0,125	0,25
00649527	0	0
05983092	0,25	0
00165968	0,125	0,25
03753293	0	0
11649393	0	0

Hasil sentimen	0,75	0,75
Hasil data	Positif	

Dari hasil yang kita proses diatas kita dapat melihat dari data awal yang memiliki opini negatif pada perhitungan manualnya dan ketika dijalankan dalam program memiliki hasil opini yang positif. Dan terkadang hasil analisa kurang maksimal karena beberapa faktor penyebabnya, dibawah ini adalah faktor penyebab kurang maksimalnya hasil analisa. Kesalahan dalam mendekteksi data

Kesalahan ini bisa terjadi karena dalam penghitungan manual kita memberikan nilai positif pada suatu kalimat namu setelah di proses oleh sistem hasilnya menjadi negatif. Penyebab kesalahan ini sering terjadi adalah munculnya kata “no atau not(n’t)” pada suatu kalimat, kata tersebut memiliki nilai positif yang tinggi didalam proses tersebut. contohnya kalimat ini “*i still remember my password, says its invalid, what?! even using the "forget password", i didn't get the email. i entered my valid email address and i still didn't receive it*”. Kalimat ini secara manual diasumsikan sebagai data positif namun hasil dalam program setelah diproses adalah negatif.

4. PENUTUP

Tujuan penelitian ini dilakukan ialah mengurangi kemunculan ambiguitas yang memiliki beberpaa makna dengan cara memanfaatkan algoritma similaritas *Wu & Palmer*, sehingga dapat mengurangi kesalaham yang terjadi dalam menafsirkan sebuah kalimat.

Dari pembahasan yang dilakukan diatas maka dapat diambil kesimpulan sebagai berikut.

4.1.Masalah dari ambiguitas kalimat belum teratasi sepenuhnya, namun dari kalimat-kalimat yang diujikan dalam sistem, tidak semua kalimat akan mendapatkan sysnset yang ditentukan oleh system dan kesalahan system dalam pendeteksian synset.

4.2.Dalam penggunaan metode *Wu & Palmer* jelas sudah terjadi peningkatan performa yang lebih baik dibandingkan metode *first sense* dan *average*. Hal ini dapat dilihat pada perbandingan *precision*, *recall*, dan *accuracy* pada Tabel 2 tadi. Disitu memperlihatkan perbandingan yang jelas, dari metode *fisrt sense* dan *average* yang menunjukkan bahwa nilai dari kedua metode tersebut masih berada dibawah dari pada nilai menggunakan

metode *Wu & Palmer*. Dapat dilihat juga hasil dari nilai *precision*, *recall*, dan *average*. Untuk *precision* menggunakan metode *Wu & Palmer* mengalami peningkatan sebesar 0,06 sampai 0,08. Kemudian *recall* mengalami peningkatan sebesar 0,22 sampai 0,24. Yang terakhir adalah *accuracy* yang mengalami peningkatan 0,04 sampai 0,06.

4.3. Jika semakin banyak data yang digunakan maka nantinya hasil klasifikasi juga akan semakin baik. Dari data yang banyak tersebut akan mempengaruhi juga dalam mengukur performa dalam mengkategorikan hasil *precision*, *recall*, dan *average*.

DATAR PUSTAKA

- Denecke, K. (2008). Using Sentiwordnet for Multilingual Sentiment Analysis. *IEEE 24th International Conference on Data Engineering Workshop*.
- Kapadia, D., and Jufrizal (2013). Types of Semantic Ambiguity Found in the Editorials of Jakarta Post Daily Newspaper. *Padang: Universitas Negri Padang*.
- Kusumawati, I., & Pamungkas, E. W. (2017). Analisa Sentimen Menggunakan Lexicon Based Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Rokok Pada Media Sosial Twitter (*Doctoral dissertation, Universitas Muhammadiyah Surakarta*).
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures On Human Language Technologies*, 5(1), 1-167.
- Pamungkas, E. W., & Putri, D. P. (2016). An Experimental Study of Lexicon-Based Sentiment Analysis on Bahasa Indonesia. *Proceeding of The 6th International Annual Engineering Seminar (INAES)*.
- Vinodhini, G. And RM. Chandrasekaran (2012). "Sentiment Analysis and Opinion Mining: A Survey". *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol.2, No.6, pp.283-292.
- Wicaksana, I. W. S., & Reza A. H. (2006). Pendekatan Schema Matching dalam Bahasa Indonesia. *Depok: Universitas Gunadarma*.